MEME16203 Linear Models Marking Guide

Test 1

UNIVERSITI TUNKU ABDUL RAHMAN

Faculty:	FES	Unit Code:	MEME16203
Course:	MAC	Unit Title:	Linear Models
Year:	1,2	Lecturer:	Dr Yong Chin Khian
Session:	202205		

Show your workings. If no workings are shown, ZERO is awarded.

Q1. Suppose $\mathbf{X}$ and $\mathbf{W}$ are any two matrices with n rows for which $C(\mathbf{X}) = C(\mathbf{W})$. Define $\mathbf{P_X} = \mathbf{X}(\mathbf{X^TX})^{-}\mathbf{X^T}$ and $\mathbf{P_W} = \mathbf{W}(\mathbf{W^TW})^{-}\mathbf{W^T}$. Show that $\mathbf{P_XP_W} = \mathbf{P_W}$. (10 marks)

Ans.

$C(\mathbf{X}) = C(\mathbf{W})$
$\Rightarrow \mathbf{W} = \mathbf{XF}$ for some $\mathbf{F}$ and $\mathbf{X} = \mathbf{WG}$ for some $\mathbf{G}$. Thus,

$$\begin{aligned}\mathbf{P_XP_W} &= \mathbf{P_XW}(\mathbf{W^TW})^{-}\mathbf{W^T}\\ &= \mathbf{P_XXF}(\mathbf{W^TW})^{-}\mathbf{W^T}\\ &= \mathbf{XF}(\mathbf{W^TW})^{-}\mathbf{W^T}\\ &= \mathbf{W}(\mathbf{W^TW})^{-}\mathbf{W^T}\\ &= \mathbf{P_W}\end{aligned}$$

Q2. Suppose $\mathbf{W}$ is an $n \times k$ matrix and $\mathbf{D}$ is a $k \times k$ non-singular matrix. Define $\mathbf{P_W} = \mathbf{W}(\mathbf{W^TW})^{-}\mathbf{W^T}$, prove that $\mathcal{C}(\mathbf{P_W}) = \mathcal{C}(\mathbf{P_WD^{-1}})$. (15 marks)

Ans.

Prove that $\mathcal{C}(\mathbf{P_W}) = \mathcal{C}(\mathbf{P_WD^{-1}})$:

$$\begin{aligned}\mathbf{a} \in \mathcal{C}(\mathbf{P_W}) &\iff \mathbf{a} = \mathbf{P_Wb} \text{ for some } \mathbf{b}\\ &\iff \mathbf{a} = \mathbf{P_WIb} \text{ for some } \mathbf{b}\\ &\iff \mathbf{a} = \mathbf{P_WD^{-1}}\underbrace{\mathbf{Db}}_{p\times 1} \text{ treat as } \mathbf{P_WD^{-1}} \text{ product a } k \times 1 \text{ vector}\\ &\Longrightarrow \mathbf{a} \in \mathcal{C}(\mathbf{P_WD^{-1}})\end{aligned}$$

So, $\mathcal{C}(\mathbf{P_W}) \subseteq \mathcal{C}(\mathbf{P_WD^{-1}})$
Then similarly,

$$\begin{aligned}\mathbf{g} \in \mathcal{C}(\mathbf{P_WD^{-1}}) &\iff \mathbf{g} = \mathbf{P_WD^{-1}h} \text{ for some } \mathbf{h}\\ &\iff \mathbf{g} = \mathbf{P_W}\underbrace{\mathbf{D^{-1}h}}_{p\times 1} \text{ treat as } \mathbf{P_W} \text{ product a } p \times 1 \text{ vector}\\ &\Longrightarrow \mathbf{g} \in \mathcal{C}(\mathbf{P_W})\end{aligned}$$

So, $\mathcal{C}(\mathbf{P_W D^{-1}}) \subseteq \mathcal{C}(\mathbf{P_W})$
And hence,
$\mathcal{C}(\mathbf{P_W}) = \mathcal{C}(\mathbf{P_W D^{-1}})$

Q3. A useful result from linear algebra (that you may use it without proof) is as follows:

$$\text{rank}(\mathbf{AB}) \leq \min[\text{rank}(\mathbf{A}), \text{rank}(\mathbf{B})]$$

for any two matrices $\mathbf{A}$ and $\mathbf{B}$ with dimensions that allow multiplication (number of columns of $\mathbf{A}$ equals the number of rows of $\mathbf{B}$). In words, this result says that the rank of a product of matrices is no greater than the rank of any matrix in the product. Show that f $\mathbf{X}$ is an $n \times p$ matrix and $\mathbf{W}$ is a matrix with n columns satisfying $\mathbf{WP_X} = \mathbf{W}$, then $\text{rank}(\mathbf{WX}) = \text{rank}(\mathbf{W})$. (15 marks)

Ans.

We are given that for any two matrices $\mathbf{A}$ and $\mathbf{B}$ that allow for the product matrix $\mathbf{AB}$,

$$\text{rank}(\mathbf{AB}) \leq \min[\text{rank}(\mathbf{A}), \text{rank}(\mathbf{B})]$$

This says rank(AB) is no larger than the smaller of the two quantities $\text{rank}(\mathbf{A})$ and $\text{rank}(\mathbf{B})$,which implies

$$\text{rank}(\mathbf{AB}) \leq \text{rank}(\mathbf{A}) \text{ and } \text{rank}(\mathbf{AB}) \leq \text{rank}(\mathbf{B}).$$

(i) $\text{rank}(\mathbf{WX}) \leq \text{rank}\mathbf{W}$
(ii)

$$\begin{aligned}
\text{rank}(\mathbf{W}) &= \text{rank}(\mathbf{WP_X}) \text{ since } \mathbf{WP_X} = \mathbf{W} \\
&\leq \text{rank}(\mathbf{WX(X^TX)^-X^T}) \\
&\leq \text{rank}(\mathbf{WX})
\end{aligned}$$

$\therefore \text{rank}(\mathbf{WX}) = \text{rank}(\mathbf{P_W})$.

Q4. Suppose that we are interested in the coefficients $\boldsymbol{\beta}$ of a linear model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, where $\mathbf{Y}$ is $n \times 1$, $\mathbf{X}$ is nonsingular with dimension $n \times p$ and $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \sigma^2\mathbf{I})$. Furthermore, suppose that it is of interest to partition that model in the form $\mathbf{X} = \begin{bmatrix}\mathbf{X}_1 & \mathbf{X}_2\end{bmatrix}$, for $n \times p_i$ matrices $\mathbf{X}_i, i = 1, 2$. Finally, suppose that an investigator creates a partially orthogonal design, in which $\mathbf{X} = \begin{bmatrix}\mathbf{X}_1 & \mathbf{X}_2\end{bmatrix}$ has the property that $\mathbf{X_1^T X_2} = 0$. You are given that the least squares estimate of $\boldsymbol{\beta}$ takes the form $\hat{\boldsymbol{\beta}} = \begin{bmatrix}\hat{\beta}_1 \\ \hat{\beta}_2\end{bmatrix}$, where

- $\hat{\boldsymbol{\beta}}_\mathbf{1} = (\mathbf{X_1^T X_1})^{-1}\mathbf{X_1^T Y}$

- $\hat{\boldsymbol{\beta}}_2 = (\mathbf{X_2^T X_2})^{-1}\mathbf{X_2^T Y}$

Show that the estimates $\hat{\boldsymbol{\beta}}_1$ and $\hat{\boldsymbol{\beta}}_2$ are uncorrelated. (15 marks)

Ans.

$$
\begin{aligned}
V(\hat{\boldsymbol{\beta}}) &= V\begin{bmatrix}(\mathbf{X_1^T X_1})^{-1}\mathbf{X_1^T}\\ (\mathbf{X_2^T X_2})^{-1}\mathbf{X_2^T}\end{bmatrix}\mathbf{Y}\\
&= \begin{bmatrix}(\mathbf{X_1^T X_1})^{-1}\mathbf{X_1^T}\\ (\mathbf{X_2^T X_2})^{-1}\mathbf{X_2^T}\end{bmatrix} V(\mathbf{Y}) \left[\mathbf{X_1}(\mathbf{X_1^T X_1})^{-1} \ \mathbf{X_2}(\mathbf{X_2^T X_2})^{-1}\right]\\
&= \begin{bmatrix}(\mathbf{X_1^T X_1})^{-1}\mathbf{X_1^T}\\ (\mathbf{X_2^T X_2})^{-1}\mathbf{X_2^T}\end{bmatrix} \sigma^2(\mathbf{I}) \left[\mathbf{X_1}(\mathbf{X_1^T X_1})^{-1} \ \mathbf{X_2}(\mathbf{X_2^T X_2})^{-1}\right]\\
&= \begin{bmatrix}(\mathbf{X_1^T X_1})^{-1}\mathbf{X_1^T X_1}(\mathbf{X_1^T X_1})^{-1} & (\mathbf{X_1^T X_1})^{-1}\mathbf{X_1^T X_2}(\mathbf{X_2^T X_2})^{-1}\\ (\mathbf{X_2^T X_2})^{-1}\mathbf{X_2^T X_1}(\mathbf{X_1^T X_1})^{-1} & (\mathbf{X_2^T X_2})^{-1}\mathbf{X_2^T X_2}(\mathbf{X_2^T X_2})^{-1}\end{bmatrix}\\
&= \begin{bmatrix}(\mathbf{X_1^T X_1})^{-1} & \mathbf{0}\\ \mathbf{0} & (\mathbf{X_2^T X_2})^{-1}\end{bmatrix}
\end{aligned}
$$

Q5. Suppose that y_{11} and y_{12} are independent $N(\mu_1, 4\sigma^2)$ variables independent of y_{21} and y_{22} that are independent $N(\mu_2, 16\sigma^2)$ and $N(\mu_2, 9\sigma^2)$ variables respectively. What is the BLUE of $6\mu_1 + 2\mu_2$? Explain carefully. (15 marks)

Ans.

$$
\begin{bmatrix}y_{11}\\ y_{12}\\ y_{21}\\ y_{22}\end{bmatrix} = \begin{bmatrix}1 & 0\\ 1 & 0\\ 0 & 1\\ 0 & 1\end{bmatrix}\begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix} + \boldsymbol{\epsilon}
$$

$$
\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}
$$

where

$\boldsymbol{\epsilon} \sim N(\mathbf{0}, \sigma^2 diag(4, 4, 16, 9))$

$$
\mathbf{X} = \begin{bmatrix}1 & 0\\ 1 & 0\\ 0 & 1\\ 0 & 1\end{bmatrix} \quad \text{and} \quad \mathbf{Y} = \begin{bmatrix}y_{11}\\ y_{12}\\ y_{21}\\ y_{22}\end{bmatrix}
$$

Let $\mathbf{V^{-1/2}} = diag(\frac{1}{\sqrt{4}}, \frac{1}{\sqrt{4}}, \frac{1}{\sqrt{16}}, \frac{1}{\sqrt{9}})$ and

$$
\mathbf{Z} = \mathbf{V^{-1/2}Y} = \begin{bmatrix}\frac{1}{\sqrt{4}}y_{11}\\ \frac{1}{\sqrt{4}}y_{12}\\ \frac{1}{\sqrt{16}}y_{21}\\ \frac{1}{\sqrt{9}}y_{22}\end{bmatrix}
$$

$$\begin{aligned} E(\mathbf{Z}) &= \mathbf{V^{-1/2}}E(\mathbf{Y}) \\ &= \mathbf{V^{-1/2}X}\boldsymbol{\beta} \\ &= \mathbf{W}\boldsymbol{\beta} \end{aligned}$$

where

$$\mathbf{W} = \begin{bmatrix} \frac{1}{\sqrt{4}} & 0 \\ \frac{1}{\sqrt{4}} & 0 \\ 0 & \frac{1}{\sqrt{16}} \\ 0 & \frac{1}{\sqrt{9}} \end{bmatrix}$$

$$\begin{aligned} Var(\mathbf{Z}) &= \mathbf{V^{-1/2}}^T Var(\mathbf{Y})\mathbf{V^{-1/2}} \\ &= \begin{bmatrix} \frac{1}{\sqrt{4}} & & & \\ & \frac{1}{\sqrt{4}} & & \\ & & \frac{1}{\sqrt{16}} & \\ & & & \frac{1}{\sqrt{9}} \end{bmatrix} \begin{bmatrix} 4\sigma^2 & & & \\ & 4\sigma^2 & & \\ & & 16\sigma^2 & \\ & & & 9\sigma^2 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{4}} & & & \\ & \frac{1}{\sqrt{4}} & & \\ & & \frac{1}{\sqrt{16}} & \\ & & & \frac{1}{\sqrt{9}} \end{bmatrix} \\ &= \sigma^2\mathbf{I} \end{aligned}$$

Thus, $\mathbf{Z}$ follows a Gauss-Markov model with model matrix $\mathbf{W}$.

$$(\mathbf{W^TW}) = \begin{bmatrix} \frac{1}{\sqrt{4}} & \frac{1}{\sqrt{4}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{16}} & \frac{1}{\sqrt{9}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{4}} & 0 \\ \frac{1}{\sqrt{4}} & 0 \\ 0 & \frac{1}{\sqrt{16}} \\ 0 & \frac{1}{\sqrt{9}} \end{bmatrix} = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.1736111111111111 \end{bmatrix}$$

$$(\mathbf{W^TW})^{-1} = \begin{bmatrix} 2.0 & 0 \\ 0 & 5.76 \end{bmatrix}$$

$$\mathbf{W^TZ} = \begin{bmatrix} \frac{1}{\sqrt{4}} & \frac{1}{\sqrt{4}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{16}} & \frac{1}{\sqrt{9}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{4}}y_{11} \\ \frac{1}{\sqrt{4}}y_{12} \\ \frac{1}{\sqrt{16}}y_{21} \\ \frac{1}{\sqrt{9}}y_{22} \end{bmatrix} = \begin{bmatrix} \frac{1}{4}y_{11} + \frac{1}{4}y_{12} \\ \frac{1}{16}y_{21} + \frac{1}{9}y_{22} \end{bmatrix}$$

So, the BLUE of $6\mu_1 - 2\mu_2$ is

$[6 \quad 2][(\mathbf{W^TW})^{-1}\mathbf{W^TZ}]$

$= [6 \quad 2] \begin{bmatrix} 2.0 & 0 \\ 0 & 5.76 \end{bmatrix} \begin{bmatrix} \frac{1}{4}y_{11} + \frac{1}{4}y_{12} \\ \frac{1}{16}y_{21} + \frac{1}{9}y_{22} \end{bmatrix}$

$= 12.0[\frac{1}{4}y_{11} + \frac{1}{4}y_{12}] + 11.52[\frac{1}{16}y_{21} + \frac{1}{9}y_{22}]$

MEME16203 Linear Models Marking Guide

Q6. You are given:

- Model (1): $Y_{ij} = \mu + \alpha_i + \epsilon_{ij}$
- Model (2): $Y_{ij} = \gamma_0 + \gamma_1(X_i - 4) + \gamma_2(X_i - 4)^2 + \epsilon_{ij}$

where $i = 1, 2, 3,\ j = 1, 2,\ X_1 = 2, X_2 = 4, X_3 = 6$ and μ, α_1, α_2, α_3, γ_0, γ_1, γ_2 and γ_3 are unknown parameters.

(a) For model (2), write down a formula for the best linear unbiased estimator (BLUE) for $\boldsymbol{\gamma} = \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \end{bmatrix}$. (10 marks)

Ans.

$$\mathbf{Y} = \begin{bmatrix} Y_{11} \\ Y_{12} \\ Y_{21} \\ Y_{22} \\ Y_{31} \\ Y_{32} \end{bmatrix},$$

$$\mathbf{X} = \begin{bmatrix} 1 & -2 & 4 \\ 1 & -2 & 4 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 2 & 4 \\ 1 & 2 & 4 \end{bmatrix}, \mathbf{X^T X} = \begin{bmatrix} 6 & 0 & 16 \\ 0 & 16 & 0 \\ 16 & 0 & 64 \end{bmatrix} \text{ and } \mathbf{X^T Y} = \begin{bmatrix} Y_{..} \\ 2(Y_{3.} - Y_{1.}) \\ 4(Y_{1.} + Y_{3.}) \end{bmatrix}$$

$$|\mathbf{X^T X}| = 6(16)(64) - 16^3 = 2048$$

$$(\mathbf{X^T X})^{-1}| = \frac{1}{2048} \begin{bmatrix} \begin{vmatrix} 16 & 0 \\ 0 & 64 \end{vmatrix} & \begin{vmatrix} 0 & 0 \\ 16 & 64 \end{vmatrix} & \begin{vmatrix} 0 & 16 \\ 16 & 0 \end{vmatrix} \\ \begin{vmatrix} 0 & 16 \\ 0 & 64 \end{vmatrix} & \begin{vmatrix} 6 & 16 \\ 16 & 64 \end{vmatrix} & \begin{vmatrix} 6 & 0 \\ 16 & 0 \end{vmatrix} \\ \begin{vmatrix} 0 & 16 \\ 16 & 0 \end{vmatrix} & \begin{vmatrix} 6 & 16 \\ 0 & 0 \end{vmatrix} & \begin{vmatrix} 6 & 0 \\ 0 & 16 \end{vmatrix} \end{bmatrix} =$$

$$\begin{bmatrix} 0.5 & 0 & -0.125 \\ 0 & 0.0625 & 0 \\ -0.125 & 0 & 0.046875 \end{bmatrix}$$

$$\hat{\gamma} = (\mathbf{X^T X})^{\mathbf{1}} \mathbf{X^T Y} = \begin{bmatrix} 0.5 & 0 & -0.125 \\ 0 & 0.0625 & 0 \\ -0.125 & 0 & 0.046875 \end{bmatrix} \begin{bmatrix} Y_{..} \\ 2(Y_{3.} - Y_{1.}) \\ 4(Y_{1.} + Y_{3.}) \end{bmatrix} =$$

$$\begin{bmatrix} \bar{Y}_{2.} \\ \frac{\bar{Y}_{3.} - \bar{Y}_{1.}}{4} \\ \frac{\bar{Y}_{1.} - 2\bar{Y}_{2.} + \bar{Y}_{3.}}{8} \end{bmatrix}$$

(b) For model (1), verify that $\tau = 5\mu + 2\alpha_1 - 4\alpha_2 + 7\alpha_3$ is an estimable function and writhe down a formula for $\hat{\tau}$, the BLUE for τ. (10 marks)

Ans.

$$\begin{aligned}\tau &= 5\mu + 2\alpha_1 - 4\alpha_2 + 7\alpha_3\\ &= 2(\mu+\alpha_1) - 4(\mu+\alpha_2) + 7(\mu+\alpha_3)\\ &= E(2\bar{Y}_{1.} - 4\bar{Y}_{2.1} + 7\bar{Y}_{3.})\end{aligned}$$

Hence, τ is estimable.
The BLUE for τ is $2\bar{Y}_{1.} - 4\bar{Y}_{2.1} + 7\bar{Y}_{3.}$.

(c) Formulate what is meant by the statement that model (2) is a reparameterization of model (1), and verify that this statement is correct. (10 marks)

Ans.
The model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ is a reparameterization of the the model $\mathbf{Y} = \mathbf{W}\boldsymbol{\gamma} + \boldsymbol{\epsilon}$ if there is a matrix $\mathbf{F}$ such that $\mathbf{W} = \mathbf{XF}$ and a matrix $\mathbf{G}$ such that $\mathbf{X} = \mathbf{WG}$. The the space spanned by the columns of $\mathbf{X}$ is a basis for space spanned by the columns of $\mathbf{W}$ and vice versa.

In this case we can write model (1) as

$$\mathbf{Y} = \mathbf{X}\begin{bmatrix}\mu\\ \alpha_1\\ \alpha_2\\ \alpha_3\end{bmatrix} + \boldsymbol{\epsilon},$$

where

$$\mathbf{X} = \begin{bmatrix}1&1&0&0\\1&1&0&0\\1&0&1&0\\1&0&1&0\\1&0&0&1\\1&0&0&1\end{bmatrix},$$

and we can write model (2) as

$$\mathbf{Y} = \mathbf{W}\begin{bmatrix}\gamma_0\\ \gamma_1\\ \gamma_2\end{bmatrix} + \boldsymbol{\epsilon},$$

where

$$\mathbf{W} = \begin{bmatrix}1&-2&4\\1&-2&4\\1&0&0\\1&0&0\\1&2&4\\1&-2&4\end{bmatrix}$$

Then,

$$\mathbf{W} = \mathbf{X}\mathbf{F}, \text{ where } \mathbf{F} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 4 \\ 0 & 0 & 0 \\ 0 & 2 & 4 \end{bmatrix}, \text{ and}$$

$$\mathbf{X} = \mathbf{W}\mathbf{G} \text{ where } \mathbf{G} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & -\frac{1}{4} & 0 & \frac{1}{4} \\ 0 & \frac{1}{8} & -\frac{1}{4} & \frac{1}{8} \end{bmatrix}.$$